

Entropy Theory for Cross Sections

Nir Avni

November 10, 2009

Abstract

We define the notion of entropy for a cross-section of an action of continuous amenable group, and relate it to the entropy of the ambient action. As a result, we are able to answer a question of J.P. Thouvenot about completely positive entropy actions.

1 Introduction

1.1 Cross Sections of Flows

One of the oldest techniques for studying continuous flows is to restrict them to cross-sections. For a smooth flow—that is, an action of the group \mathbb{R} by diffeomorphisms on a smooth manifold M —Poincare defined a cross-section as a co-dimension 1 sub-manifold, N , that is transverse to the direction of the flow. The transversality condition implies that each orbit intersects N in a discrete set. In other words, for each point x of M , the set of visit times, $\{t \in \mathbb{R} | T_t(x) \in N\}$ is discrete. The co-dimension 1 condition implies that N intersect many orbits. The Poincare return map assigns to a point $x \in N$ the first point in the positive trajectory of x that lies in N .

Extending this construction, Ambrose proved (see [Am]) that for any Borel flow on a Borel probability space (X, μ) , there is a Borel subset $Y \subset X$ such that, for almost all points $x \in X$, the set $\{t \in \mathbb{R}_{>0} | T_t(x) \in Y\}$ is non-empty and discrete. If the flow preserves the measure μ , then there is a measure ν on Y such that, near Y , the measure μ is a product of ν and the Lebesgue measure (in the direction of the flow). It follows that for ν -almost all $y \in Y$, the set $\{t \in \mathbb{R}_{>0} | T_t(y) \in Y\}$ is also nonempty and discrete. Therefore, this set has

a minimum, which we call $\alpha(y)$. Define a transformation $S : N \rightarrow N$ by $S(y) = T_{\alpha(y)}(y)$. One can show that S preserves the measure ν .

In this way, a flow (X, T_t) induces a measure preserving transformation (Y, S) . Conversely, Ambrose shows that the flow can be obtained from the triple (Y, S, α) by the construction of the ‘flow under a function’. There are many connections between the dynamics of the flow T_t and the dynamics of the transformation S ; one of the basic ones is the Abramov formula (see [CFS]) connecting the entropies of the two actions:

$$h(S) = h(T_1) \int_Y \alpha(y) d\nu. \quad (1)$$

The method of cross-sections is less-understood for higher dimensional groups. In [FHM], the authors show that every Borel action of a locally-compact and second-countable group admits a cross-section (see Definition 2.2 below). However, even for the group \mathbb{R}^2 , the cross-section does not come equipped with an action of a discrete subgroup of the continuous group. In this case, it is not clear what is the higher dimensional analogue of Abramov formula.

1.2 Main Results

The purpose of this article is to develop the technique of cross-sections for probability-preserving free actions of a class of amenable groups. This class contains, for example, all nilpotent Lie groups. In particular, we state and prove a generalization of Abramov formula for these groups. As a corollary, we deduce a generalization of a theorem of Rohlin and Sinai on Kolmogorov systems. Before stating the theorem, we give some definitions. Let (X, \mathcal{B}, μ) be a probability space. For a partition P of X , whose parts have measures p_1, \dots, p_n , let $H(P) = \sum p_i \log p_i$. A transformation $T : X \rightarrow X$ is called *uniformly mixing* if, for every partition P and $\epsilon > 0$, there is an integer N such that for every k and every finite sequence of integers $i_1 < i_2 < \dots < i_k$ satisfying $i_{j+1} - i_j \geq N$ for all j , we have

$$\left| \frac{1}{k} H(T^{i_1} P \vee T^{i_2} P \vee \dots \vee T^{i_k} P) - h(T, P) \right| < \epsilon.$$

Here $h(T, P)$ is the (measure theoretic) entropy of T with respect to the partition P .

Recall that the *spectrum* of (X, \mathcal{B}, m, T) is defined as the spectrum of the unitary operator $U_T : L_2(X, \mathcal{B}, m) \rightarrow L_2(X, \mathcal{B}, m)$ given by $(U_T f)(x) = f(Tx)$.

In [RS], the following theorem is proved:

Theorem 1.1. *Let (X, \mathcal{B}, m) be a probability space and let $T : X \rightarrow X$ be a probability-preserving transformation. Assume that, for any nontrivial partition P of X , we have $h(T, P) > 0$. Then*

1. *The transformation T is uniformly mixing.*
2. *The spectrum of (X, \mathcal{B}, m, T) is Lebesgue with countable multiplicity.*

An entropy theory for actions of amenable groups is developed in [OW]. The definitions there do not apply to all amenable groups, but only to a class of amenable groups called *groups with good entropy theory*. Let G be an amenable group with good entropy theory. Given a probability-preserving action of G on a probability space X and a partition P of X , the entropy of the action with respect to the partition will be denoted by $h(G, P)$.

The entropy of an action relative to a sub-sigma-algebra is defined only for a smaller class of amenable groups. These are called in [OW] *groups with zero self entropy*. We note here that all nilpotent Lie groups have zero self entropy, and that all groups with zero self entropy (and, in fact, all groups with good entropy theory) are unimodular. For more information on groups with zero self entropy, see Section 3. Starting from Section 3, we shall assume that the groups we talk about have zero self entropy.

For technical reasons, we deal with Borel actions only. This means that the probability space X is a Polish space endowed with the Borel sigma-algebra, that the probability measure is regular, and that the action map $G \times X \rightarrow X$ is Borel measurable. In fact, as remarked in [FHM], since we are interested only in measure theoretic properties, we can change X and the action to an isomorphic (in the category of measure-preserving actions) Borel action, which is also continuous.

An action of G on X is called *completely positive entropy* (or CPE for short) if, for any non-trivial partition P of X , the entropy $h(G, P)$ is strictly positive. It is called *free* if, for every $g \in G$, which is different from 1, the set of fixed points of g is negligible.

Definition 1.2. Let G be a group and let $K \subset G$ be a subset. A finite subset $F \subset G$ is called *K-separated* if, for every two non-equal elements g, h of F , the element gh^{-1} does not belong to K .

This is the generalization of Theorem 1.1:

Theorem 1.3. *Let G be an amenable group with zero self entropy. Suppose G acts freely on a probability space X , and suppose that the action is CPE. Then*

1. *For every partition P and any $\epsilon > 0$, there is a compact set $K \subset G$ such that for any finite set $F \subset G$ that is K -separated,*

$$\left| \frac{1}{|F|} H \left(\bigvee_{g \in F} gP \right) - H(P) \right| < \epsilon.$$

2. *As a G -module, the space $L_2(X)$ decomposes as a direct sum of infinitely many copies of the regular representation of G .*

This theorem was conjectured by Thouvenot. It was proved for discrete amenable groups in [RW] and [DG]; see also [D] and [DP].

1.3 Organization

In the rest of this introduction, we shall outline, without proofs, the main points of this paper.

Let G be a locally compact, second countable, and unimodular group. Let X be a Polish space, let \mathcal{B} be the Borel sigma-algebra on X , and let m be a Borel probability measure on X . For a free, Borel, and measure-preserving action of G on (X, \mathcal{B}, m) , we define a *cross-section* to be a Borel set $S \subset X$ that intersects almost every orbit in a discrete set. On S , we have the Borel sigma-algebra \mathcal{B}_S . We shall show that there is a canonically defined Borel measure μ on S , such that, locally near S , the measure m is the product of μ and the Haar measure of G (in the direction of the action).

In contrast to the one-dimensional case, S does not come equipped with a canonical action of a discrete group. There are, however, additional structures on S . First, there is an equivalence relation: $x \sim y$ if and only if x and y (both are elements of S) lie in the same G orbit. Denote by $\mathcal{R} \subset S \times S$ the set of equivalent pairs. We also get a function $\alpha : \mathcal{R} \rightarrow G$ by defining $\alpha(x, y) = g$ if $gy = x$ (this g is unique because the action is assumed to be free). Note that we used the same letter, α , as before, since it is an analogue of the function we introduced in the one-dimensional case.

We concentrate on quintuples of the form $(S, \mathcal{B}, \mu, \mathcal{R}, \alpha)$. Note that this information contains G implicitly. Given $\mathfrak{S} = (S, \mathcal{B}, \mu, \mathcal{R}, \alpha)$, it is

possible to formulate an analogue of the mean ergodic theorem. If \mathfrak{S} comes from a cross-section of an action of G , then this analogue of the mean ergodic theorem holds. Surprisingly, the converse is also true: the mean ergodic theorem for \mathfrak{S} implies that \mathfrak{S} is isomorphic to a cross-section of a probability-preserving action. These constructions are described in Section 2. In the rest of the article, we only deal with such quintuples, which we call cross-sections (slightly abusing notations). In Section 2, we also describe our main technical tools—a tiling lemma that generalizes [OW, Proposition 7], and an ergodic theorem for (generalizations of) Rohlin towers.

Section 3 is devoted to the entropy theory of cross-sections. Building on [OW], we define entropy for cross-sections of actions of groups with zero self entropy, and prove an analogue of Abramov’s theorem. Interestingly, the useful notion here is relative entropy (with respect to a sub-sigma-algebra) rather than the absolute entropy.

In the same section, we prove the following transfer theorem (see also [RW]): Let G, H be two amenable groups with zero self entropy, let $\mathfrak{S} = (S, \mathcal{B}, \mu, \mathcal{R}, \alpha)$ be a cross-section for an action of G , and let $\mathfrak{S}' = (S', \mathcal{B}', \mu', \mathcal{R}', \beta)$ be a cross-section for an action of H . Suppose that $\phi : S \rightarrow S'$ is a measure-preserving and equivalence-preserving map (i.e. $x\mathcal{R}y \iff \phi(x)\mathcal{R}'\phi(y)$), and assume that $\beta \circ \phi$ is measurable with respect to a sub-sigma-algebra $\mathcal{G} \subset \mathcal{B}$. Then, the relative entropies of \mathfrak{S} and \mathfrak{S}' with respect to \mathcal{G} and $\phi_*\mathcal{G}$ are the same.

This transfer theorem is useful to the ergodic theorem of groups. By a theorem of [CFW], for any \mathfrak{S} and \mathfrak{S}' as above, there is an isomorphism ϕ as above (in general, however, we cannot say much about \mathcal{G}). This allows us to transfer questions (and answers) from actions of one group to actions of another, using cross-sections as intermediaries.

In Section 4, we prove Theorem 1.3. Section 5 features some concluding remarks.

1.4 Acknowledgment

This work contains results of a research done under the supervision of Prof. Benjamin Weiss at the Hebrew University. I thank him for being such a good source of interesting questions, for so many illuminating discussions, and for his help in the process of writing this paper. I also thank Prof. Dan Rudolph, Prof. Valentyn Golodets, and the referees for helpful remarks.

2 Ergodic Theory for Cross Sections

2.1 Cross Sections

Convention 2.1. Unless stated otherwise, all groups will be locally compact, second countable, unimodular, and amenable. Starting from the next section, we also assume that they have zero self entropy. All spaces are assumed to be Borel, all sigma-algebras are assumed to be contained in the Borel sigma-algebra, and all actions are assumed to be continuous, probability-preserving, and free.

Suppose G is a group as above with Haar measure λ . Recall that G is called *amenable* if, for every $\epsilon > 0$ and a compact $K \subset G$, there is a compact $F \subset G$ such that $\lambda(KFK) < (1+\epsilon)\lambda(F)$. Such an F is called (K, ϵ) -invariant. We say that a compact set F is *sufficiently invariant* if F is (K, ϵ) -invariant for some K and ϵ (which should be prescribed). A sequence (F_n) of compact subsets of G is called a *Følner sequence* if, for every K and ϵ , there is an N such that F_n is (K, ϵ) -invariant for $n > N$. We refer the reader to [OW] and [W] for the ergodic theory of actions of amenable groups.

If G acts on a probability space (X, \mathcal{B}, m) , we denote the application of the group element g to the point x by $g \cdot x$.

Definition 2.2. Suppose that the group G acts freely on a probability space (X, \mathcal{B}, m) . A Borel subset $S \subset X$ is called a *cross-section* for the action if, for almost every point $x \in X$, the set of return times of x —which is defined as $\{g \in G | gx \in S\}$ —is discrete and non-empty.

By a theorem of [FHM], we know that cross-sections of free Borel actions of local compact, second countable groups always exist.

We will usually work with a slightly stronger condition. For a neighborhood $U \subset G$ of the identity, we say that the cross-section is U -discrete if, for almost any $x \in X$, the set of return times of x is U -separated. In an ergodic system, for any compact neighborhood U of the identity in G , every cross-section contains a subset which is a cross-section and is U -discrete.

Clearly, if S is a cross-section, then the restriction of the Borel sigma-algebra to S is the Borel sigma-algebra on S . We denote it by \mathcal{B} .

Definition 2.3. For a U -discrete cross-section S , the induced measure on S is

$$\mu(A) = \lim_{V \searrow \{1\}} \frac{m(V \cdot A)}{\lambda(V)}$$

where the V 's are neighborhoods of 1 in G . The limit over $V \searrow \{1\}$ means that for every open neighborhood W of 1, all but finitely many of the V 's are contained in W .

If the cross-section is U -discrete and $V^2 \subset U$, then we have a one-to-one map $a : V \times S \rightarrow X$. Pulling back the measure m on $V \cdot S$ gives a measure a^*m on $V \times S$. We decompose a^*m with respect to S : $a^*m = \int_S \eta_x d\mu(x)$. Since the action is measure-preserving, almost every η_x must be a multiple of the Haar measure λ of G . After multiplying μ by some function we can assume that $\eta_x = \lambda$ for almost every x . It is now clear that the limit above exists (it does not even depend on V if V is small enough). Also, $\mu(S) = m(V \cdot S)/\lambda(V)$, so μ is a finite measure.

Convention 2.4. By rescaling the Haar measure, we can (and will) assume $\mu(S) = 1$.

The last two ingredients we need are the equivalence relation and the cocycle induced by the action. The first is the subset $\mathcal{R} \subset S \times S$ that consists of pairs (x, y) in the same G orbit. If $(x, y) \in \mathcal{R}$ we also write $x\mathcal{R}y$. Clearly, \mathcal{R} is a Borel set and is an equivalence relation. If $(x, y) \in \mathcal{R}$, then by definition there is a $g \in G$ such that $g \cdot y = x$. This g is unique because we assumed the action was free. We set $\alpha(x, y) = g$. The function α is measurable and satisfies the equation

$$\alpha(x, y)\alpha(y, z) = \alpha(x, z).$$

Functions from an equivalence relation to a group satisfying this equation are called cocycles¹. We call α the *cocycle induced by the action*. The cocycle α that we have constructed satisfies the condition that for μ -almost all $x \in S$, the function $\alpha(x, -)$ (from the equivalence class of x to G) is one-to-one. We call cocycles that satisfy this condition *free*. Moreover, if the cross-section S is U -discrete, then for every $x, y \in S$ such that $x\mathcal{R}y$ and $x \neq y$, the element $\alpha(x, y)$ is not in U . A cocycle that satisfies this property is called U -discrete.

The following lemma is well-known.

Lemma 2.5. *Let G be a unimodular group, let $U \subset G$ be an open neighborhood of the identity, and let X be a probability space. Suppose*

¹The reader may be more familiar with the definition of cocycle of a transformation $T : X \rightarrow X$ as being a function $f : \mathbb{Z} \times X \rightarrow \mathbb{R}^\times$ that satisfies $f(n+m, x) = f(n, x) \cdot f(m, T^n x)$. This is a special case of our definition, where G is replaced by \mathbb{R}^\times , the equivalence relation consists of pairs of the form $(x, T^n x)$ and $\alpha(T^n x, x) = f(n, x)$.

that S is a U -discrete cross-section of a probability-preserving action of G on X . For every Borel transformation $T : S \rightarrow S$ such that $x\mathcal{R}(Tx)$ for almost any x , and for any $A \subset S$ such that the restriction of T to A is one-to-one, we have $\mu(TA) = \mu(A)$ (in this situation we say that \mathcal{R} preserves the measure μ).

Proof. Since $T|_A$ is one-to-one, it is enough to prove the lemma for every element in a (countable) decomposition $A = \cup A_i$ of A . Let V be a neighborhood of 1 such that $V^4 \subset U$. The inverse of the action map is a measure-preserving isomorphism $\Phi : V^2 \cdot S \rightarrow V^2 \times S$ (where on the left we take the measure m and on the right we take $\lambda \times \mu$). Consider the function $x \mapsto \alpha(x, Tx)$. There is a decomposition of A into sets A_i such that for every x, y in the same A_i ,

$$\alpha(x, Tx)\alpha(y, Ty)^{-1} \in V. \quad (2)$$

By decomposing A to the A_i , we can assume that (2) holds for every two points in A . Fix some x_0 in A , and let $g = \alpha(x_0, Tx_0)$. Let W be a neighborhood of 1 in G such that $W, g^{-1}Wg \subset V$. Since the action of G preserves m and G is unimodular,

$$\mu(A) = \frac{m(g^{-1}Wg \cdot A)}{\lambda(g^{-1}Wg)} = \frac{m(Wg \cdot A)}{\lambda(W)} = \frac{(\lambda \times \mu)(\Phi(Wg \cdot A))}{\lambda(W)}.$$

For every $x \in A$, $\Phi(Wg \cdot x)$ is of the form $Wh \times \{Tx\}$ where $h \in V$ is defined by $hTx = gx$. Hence $(\lambda \times \mu)(\Phi(Wg \cdot A)) = \lambda(W)\mu(TA)$, so $\mu(A) = \mu(TA)$. \square

2.2 The Ergodic Theorem

From now on we shall work with quintuples $(S, \mathcal{B}, \mu, \mathcal{R}, \alpha)$ where (S, \mathcal{B}) is a Borel space, μ is a probability measure on S , \mathcal{R} is a measure-preserving equivalence relation on S , and α is a free cocycle with values in an amenable group G . We stress again that G is implicitly given in this data. We wish to develop ergodic theory of such quintuples. The first step is a mean ergodic theorem:

Theorem 2.6. *Let (F_n) be a Følner sequence in G . Suppose that $(S, \mathcal{B}, \mu, \mathcal{R}, \alpha)$ is a quintuple corresponding to a cross-section of a probability-preserving free action of G , and suppose that S is U -discrete for some open neighborhood U of the identity in G . Then for any $h \in L_\infty(S, \mu)$,*

the sequence of functions

$$h_n(x) = \frac{1}{|\{(x, y) \in \mathcal{R} \wedge \alpha(x, y) \in F_n\}|} \sum_{(x, y) \in \mathcal{R} \wedge \alpha(x, y) \in F_n} h(y)$$

(and $h_n(x) = 0$ if the set $\{y | \alpha(x, y) \in F_n\}$ is empty) converges in measure to the constant function $\int_S h d\mu$.

Proof. Let M be such that $h(x) \leq M$ for almost every x . Let $V \subset G$ be a neighborhood of the identity such that $V^2 \subset U$. Define a function H on X by

$$H(x) = \begin{cases} h(s) & (\exists g \in V, s \in S) \quad x = gs \\ 0 & \text{Otherwise} \end{cases}.$$

First note that for any compact subset $F \subset G$ and for almost any $s \in S$,

$$\left| \sum_{t: \alpha(s, t) \in F} h(t) - \frac{1}{\lambda(V)} \int_F H(fs) df \right| \leq M \frac{\lambda(VF \setminus F)}{\lambda(V)}.$$

If n is big enough then F_n is (V, ϵ) -invariant, so $\lambda(VF_n \setminus F_n) < \epsilon \lambda(F_n)$. By the mean ergodic theorem for the action of G on X , we have that for any $\epsilon > 0$, if n is big enough then for all $x \in X$ outside a set of measure less than $\epsilon \lambda(V)$,

$$\left| \frac{1}{\lambda(F_n)} \int_{F_n} H(fx) df - \lambda(V) \int_S h(s) ds \right| = \left| \frac{1}{\lambda(F_n)} \int_{F_n} H(fx) df - \int_X H(y) dy \right| < \epsilon.$$

By Fubini's theorem, there is a $g \in V$ such that the above inequality holds for all $x \in gS$ except a subset of $g_*\mu$ -measure less than ϵ . Since we can assume that F_n is (V, ϵ) -invariant, we have that for *all* $x \in X$,

$$\frac{1}{\lambda(F_n)} \left| \int_{F_n} H(fx) - H(fgx) df \right| \leq \frac{2}{\lambda(F_n)} \int_{F \triangle Fg} |H(fx)| df \leq 2\epsilon M,$$

so for a subset of S of μ -measure greater than $1 - \epsilon$ we have

$$\begin{aligned} \left| \frac{1}{\lambda(F_n)} \sum_{t: \alpha(s, t) \in F_n} h(t) - \int_S h(t) dt \right| &\leq \frac{1}{\lambda(F_n) \lambda(V)} \left| \int_{F_n} H(fs) df - \int_{F_n} H(fgs) df \right| + \\ &+ \frac{(M+1)\epsilon}{\lambda(V)} \leq \frac{(2M+2)\epsilon}{\lambda(V)}. \end{aligned}$$

Taking ϵ small enough, we get that for all bounded functions h ,

$$\frac{1}{\lambda(F_n)} \sum_{t: \alpha(s,t) \in F_n} h(t) \rightarrow \int_S h(t) dt$$

in probability. Applying this to the function $h = 1$, we get that outside a set of small measure,

$$1 - \epsilon \leq \frac{|\{t : \alpha(s,t) \in F_n\}|}{\lambda(F_n)} < 1 + \epsilon.$$

Dividing the last two inequalities gives the result. \square

Remark 2.7. In a similar way, one can prove a pointwise ergodic theorem (along a tempered Følner sequence) by using [L] instead of the mean ergodic theorem. A direct argument can also be given using Proposition 2.12.

Conversely, the mean ergodic theorem implies that the quintuple comes from a G action on a probability space:

Theorem 2.8. *Suppose $(X, \mathcal{B}, \mu, \mathcal{R}, \alpha)$ is a quintuple where (X, \mathcal{B}, μ) is a probability space, \mathcal{R} is a Borel equivalence relation with countable equivalence classes that preserves μ , and $\alpha : \mathcal{R} \rightarrow G$ is a free cocycle whose range is some amenable group G . Assume that for almost all $x \in X$, the set $\{\alpha(x, y) | (x, y) \in \mathcal{R}\}$ is discrete, and that for any $f \in L_\infty(X, \mu)$, the sequence of functions f_n defined in the last theorem converges in probability to $\int f d\mu$. Then there is a probability-preserving action of G such that the quintuple is induced from this action.*

Proof. We take the Mackey range of the cocycle. That is, we look at $X \times G$ and divide by the equivalence relation $(x, g) \sim (y, \alpha(x, y)g)$ for all $(x, y) \in \mathcal{R}$ and $g \in G$. According to [FHM] the quotient $Y = X \times G / \sim$ is a Borel space.

The group G acts on Y by right multiplication. Since we assume that α is free, the map $x \mapsto [(x, 1)]$ is an isomorphism. Let $S \subset Y$ be the set $\{[(x, 1)]\}$. It follows from the assumptions that almost every G -orbit in Y intersects S in a discrete set. Since every G -orbit intersects S at least once, we get that S is a cross-section. The relation \sim on $X \times G$, as well as the right action of G preserve the measure $\mu \times \lambda$, and thus the quotient inherits a G -invariant measure. It remains to show that the mean ergodic theorem implies that this measure is finite.

Let $\pi : X \times G \rightarrow Y$ be the quotient map. We show that if $F \subset X \times G$ is Borel, its boundary has measure zero, and such that $\pi|_F$ is one-to-one, then $\mu \times \lambda(F) \leq 1$ (recall that the measure on the quotient can be computed from a Borel fundamental domain Z with negligible boundary by $m(A) = \mu \times \lambda(\pi^{-1}(A)) \cap Z$ for $A \subset X \times G / \sim$). Suppose $\mu \times \lambda(F) > 1 + 2\epsilon$. After removing from F a subset of measure less than ϵ , we can assume that there is a partition $X = X_1 \sqcup \dots \sqcup X_n$ and compact subsets K_1, \dots, K_n such that $F = \cup X_i \times K_i$. Let η be a positive small number (to be determined later), and let L be a $(\cup K_i, \eta)$ -invariant set such that all but η of the points in X satisfy the mean ergodic theorem with respect to all the characteristic functions 1_{X_i} .

Let $s \in X$ be such a point. If $t \in Ls \cap X_i$ (we shall write this as $i = i(t)$), then $K_i \alpha(s, t) \subset (\cup K_j)L$, and the sets $K_{i(t)} \alpha(s, t)$, for the different t 's, are disjoint. Therefore,

$$\begin{aligned} (1 + \eta)\lambda(L) &\geq \sum_{t \in Ls \cap X} \lambda(K_{i(t)} \alpha(s, t)) = \sum_i \lambda(K_i) |\{t | t \in Ls \cap X_i\}| \geq \\ &\geq \sum_i \lambda(K_i) \mu(X_i) \lambda(L) (1 - \eta) = (1 - \eta) \lambda(L) (\mu \times \lambda)(F), \end{aligned}$$

which is a contradiction if η is small enough. \square

Convention 2.9. From now on, by a cross-section we will mean a quintuple $(S, \mathcal{B}, \mu, \mathcal{R}, \alpha)$ —where (S, \mathcal{B}, μ) is a Borel probability space, \mathcal{R} is a Borel equivalence relation with countable equivalence classes that preserves μ , and $\alpha : \mathcal{R} \rightarrow G$ is a free cocycle which is U -discrete—that satisfies the mean ergodic theorem.

2.3 Tiling Lemma

Definition 2.10. A sequence of finite sets $A_i \subset G$ is called ϵ -disjoint if, for every i ,

$$\left| A_i \cap \bigcup_{j < i} A_j \right| < \epsilon |A_i|.$$

Lemma 2.11. Let G be an amenable group, let c be an integer, and let $0 < \delta < 0.1$. Let U and V be neighborhoods of the identity in G

such that $V^2 \subset U$, and let $F \subset G$ be a compact set such that there are at most $c\lambda(F)$ disjoint right translates of V whose centers are in F . Assume also that $\lambda(F) > 10$. Let $A \subset G$ be a finite set which is U -separated, and let $B \subset A$ be such that for every $b \in B$,

$$|Fb \cap A| > \frac{1}{2}\lambda(F).$$

Then there is a subset $\tilde{B} = \{\tilde{b}_1, \tilde{b}_2, \dots, \tilde{b}_k\} \subset B$ such that the sequence of sets $F\tilde{b}_i \cap A$, $i = 1, \dots, k$ is δ -disjoint, and such that

$$|F\tilde{B} \cap A| > \frac{\delta}{4c}|B|.$$

Proof. Let $B = \{b_1, b_2, \dots, b_{|B|}\}$. Let $\tilde{b}_1 = b_1$, and for $i > 1$ let \tilde{b}_i to be the first element of B such that

$$|F\tilde{b}_i \cap \bigcup_{j < i} F\tilde{b}_j \cap A| < \delta|F\tilde{b}_i \cap A|,$$

if such a \tilde{b} exists. Otherwise stop, and let $k = i - 1$. Let $\tilde{B} = \{\tilde{b}_1, \dots, \tilde{b}_k\}$. One of the following must hold:

1. $|\tilde{B}| > |B|/2$. In that case, every $\tilde{b}_i \in \tilde{B}$ contributes at least $\frac{1-\delta}{2}\lambda(F)$ new points to the union $F\tilde{B} \cap A$, so

$$\left| \bigcup_i F\tilde{b}_i \cap A \right| > \frac{1-\delta}{4}\lambda(F)|B| > |B|.$$

2. $|\tilde{B}| < |B|/2$. Let $E = B \setminus \tilde{B}$. For every point $b \in E$, more than a δ portion of the set $Fb \cap A$ is contained in $F\tilde{B} \cap A$. This implies that for every $b \in E$,

$$\langle 1_{Fb \cap A}, 1_{F\tilde{B} \cap A} \rangle \geq \delta|Fb \cap A|$$

where the inner product of the functions is taken with respect to counting measure. Since by our assumptions, the function $\sum_{b \in B} 1_{Fb \cap A}$ is smaller than $c\lambda(F)$,

$$c\lambda(F)|F\tilde{B} \cap A| \geq \left\langle \sum_{b \in E} 1_{Fb \cap A}, 1_{F\tilde{B} \cap A} \right\rangle \geq \delta \sum_{b \in E} |Fb \cap A| \geq \frac{1}{4}\delta\lambda(F)|B|$$

Thus, in both cases the conclusion of the lemma holds. \square

Proposition 2.12. *Let G be an amenable group, and let c be an integer. For every $0 < \delta < 0.1$ there are $\epsilon > 0$ and an integer N , such that if $U, V \subset G$ are neighborhoods of the identity, F_1, \dots, F_N are compact subsets of G , and $B \subset A \subset G$ are finite sets such that*

1. $V^2 \subset U$.
2. *For every i there are at most $c\lambda(F_i)$ disjoint translates of V whose centers are in F_i .*
3. $\lambda(F_i) > 10$.
4. A is U -separated.
5. $|B| > (1 - \epsilon)|A|$.
6. *For every i and every $b \in B$, we have*

$$\left| \left(\bigcup_{j < i} F_j \right)^{-1} F_i b \cap A \right| < (1 + \epsilon)|F_i b \cap A|.$$

Then there are subsets $\tilde{B}_i = \{\tilde{b}_{i1}, \dots, \tilde{b}_{ik_i}\} \subset B$, $i = 1 \dots N$ such that

1. *For any i , the sequence of sets $F_i \tilde{b}_{i1} \cap A, \dots, F_i \tilde{b}_{ik_i} \cap A$ is δ -disjoint.*
2. *For any $i_1 \neq i_2$, the sets $F_{i_1} \tilde{B}_{i_1} \cap A$ and $F_{i_2} \tilde{B}_{i_2} \cap A$ are disjoint.*
3. $\left| \bigcup_i F_i \tilde{B}_i \cap A \right| > (1 - 2\delta)|A|$.

Definition 2.13. We say that a sequence (A, B, F_i) satisfies the conditions of the tiling lemma with parameter ϵ if the above conditions hold.

Proof. Let $N = \frac{\log(\delta)}{\log(1-\delta/8c)}$, and let $\epsilon > 0$ be such that $(1-\delta)(1-2\epsilon) > (1-2\delta)(1+\delta+2\epsilon)$ —note that when $\epsilon = 0$, the left hand side is equal to $1-\delta$ and the right hand side is equal to $(1-2\delta)(1+\delta) = 1-\delta-2\delta^2$, which is strictly smaller than the left hand side.

We define three sequences of sets, A_i, B_i , and \tilde{B}_i for $0 < i \leq N$. Let $A_N = A$, $B_N = B$. Apply Lemma 2.11 to (A_N, B_N, F_N) to get a set \tilde{B}_N such that the collection of sets $\{F_N b \cap A_N = F_N b \cap A | b \in \tilde{B}_N\}$ is δ -disjoint, and the size of its union is bigger than $\frac{\delta}{4c}|B_N|$. Assume that A_i, B_i have been defined for $k \leq i \leq N$. If $|A_k| < 2\delta|A|$, we define $A_{k-1} = A_k, B_{k-1} = B_k, \tilde{B}_{k-1} = \emptyset$. Otherwise, let

$$A_{k-1} = A_k \setminus F_k \tilde{B}_k,$$

and

$$B_{k-1} = B_k \setminus \left(\bigcup_{j < k} F_j \right)^{-1} F_k \tilde{B}_k,$$

and apply Lemma 2.11 to $(A_{k-1}, B_{k-1}, F_{k-1})$ to get a set $\tilde{B}_{k-1} \subset B_{k-1}$, such that the collection of sets $\{F_{k-1}b \cap A_{k-1} = F_{k-1}b \cap A \mid b \in \tilde{B}_{k-1}\}$ is δ -disjoint, and the size of its union is bigger than $\frac{\delta}{4c}|B_{k-1}|$. The first two conclusions of the proposition hold. For the third, we use

Lemma 2.14. *For every k , if $|A_k| > 2\delta|A|$, then $|B_k| > |A_k|/2$.*

Proof. By decreasing induction on k . The case $k = N$ follows from assumption 5. Assuming the lemma for $k + 1, \dots, N$, let

$$\Sigma = \sum_{j=k+1}^N \sum_{b \in \tilde{B}_j} |F_j b \cap A|$$

and

$$\Sigma' = \sum_{j=k+1}^N \sum_{b \in \tilde{B}_j} \left| (\cup_{i < j} F_i)^{-1} F_j b \cap A \right|.$$

By δ -disjointness and the assumption of the lemma,

$$2\delta|A| < |A_k| \leq |A| - (1 - \delta)\Sigma \implies \Sigma < \frac{1 - 2\delta}{1 - \delta}|A|. \quad (3)$$

By assumption 6,

$$|B_k| \geq |B| - \Sigma' \geq |B| - (1 + \epsilon)\Sigma.$$

Hence, by assumption 5,

$$\frac{|B_k|}{|A_k|} \geq \frac{|B| - (1 + \epsilon)\Sigma}{|A| - (1 - \delta)\Sigma} \geq \frac{(1 - \epsilon)|A| - (1 + \epsilon)\Sigma}{|A| - (1 - \delta)\Sigma}.$$

We claim that the last expression is larger than $1/2$. This is equivalent to

$$|A|(1 - 2\epsilon) > \Sigma(1 + 2\epsilon + \delta).$$

Plugging in Inequality (3), we see that it is enough to prove that

$$(1 - 2\epsilon) > (1 + 2\epsilon + \delta) \frac{1 - 2\delta}{1 - \delta},$$

which is our assumption on ϵ . □

By the lemma, we see that if $|A_k| > 2\delta|A|$, then $|F_k\tilde{B}_k \cap A| \geq \frac{\delta}{4c}|B_k| \geq \frac{\delta}{8c}|A_k|$, which implies that $|A_{k-1}| < (1 - \frac{\delta}{8c})|A_k|$, and, using induction, that $|A_{k-1}| < (1 - \frac{\delta}{8c})^{N-k}|A|$. By our assumption on N , we get that this cannot hold for all k . Hence $|A_1| < 2\delta|A|$, which shows the third conclusion of the proposition. \square

2.4 Ergodic Theorem for Castles

We fix a cross-section $(S, \mathcal{B}, \mu, \mathcal{R}, \alpha)$.

Definition 2.15. Let $A \subset S$ be a Borel subset. A castle with base A is a Borel subset $\mathcal{T} \subset \mathcal{R}$ such that

1. The set $\{(x, x) | x \in A\}$ is contained in \mathcal{T} .
2. For all $x, y \in S$, if $(x, y) \in \mathcal{T}$ then $x \in A$.
3. For every $x \in A$, the set $\mathcal{T}_x := \{y \in S \mid (x, y) \in \mathcal{T}\}$ is finite.
4. For every two distinct points $x_1, x_2 \in A$, the sets \mathcal{T}_{x_1} and \mathcal{T}_{x_2} are disjoint.

We call \mathcal{T}_x the *tower over x* . Note that if \mathcal{T} is a castle with base A , then the function $x \mapsto \mathcal{T}_x$ from A to the finite subsets of S is Borel measurable. The *range* of the castle is the set $rg(\mathcal{T}) = \cup_{x \in A} \mathcal{T}_x$. If \mathcal{G} is a sub-sigma-algebra, we say that the castle \mathcal{T} is \mathcal{G} -measurable if \mathcal{T} is \mathcal{G} -measurable (which implies that the base of \mathcal{T} is \mathcal{G} -measurable). Finally, we say that \mathcal{T} covers δ of S , if $\mu(rg(\mathcal{T})) > \delta$.

Remark 2.16. In the Ornstein–Weiss theory, there are several notions (casle, real-tower, ϵ -quasi tower) similar to our notion of castle. The definition given here is more restrictive than ϵ -quasi tower in that we require that the sets \mathcal{T}_x are disjoint, but less restrictive in that we do not impose conditions on the possible values of $\alpha(\mathcal{T})$.

Definition 2.17. Let \mathcal{T} be a castle. Define a measure $\mu_{\mathcal{T}}$ on its base by

$$\mu_{\mathcal{T}}(B) = \int_B |\mathcal{T}_x| d\mu(x).$$

In words, $\mu_{\mathcal{T}}(B)$ is the μ -measure of the part of the range of the castle that lies over B .

Definition 2.18. Given $T \subset S$ and a compact $K \subset G$, the K -interior of T is the set

$$int_K T = \{t \in T | (\forall r \in S)(\alpha(t, r) \in K \implies r \in T)\}.$$

The K -boundary of T is $\partial_K T = T \setminus \text{int}_K T$.

Definition 2.19. Let \mathcal{T} be a castle with base A and let $K \subset G$ be compact set. For $\epsilon > 0$ we say that \mathcal{T} is (K, ϵ) -invariant if

$$\mu_{\mathcal{T}} \left(\left\{ s \in A \mid \frac{|\partial_K \mathcal{T}_s|}{|\mathcal{T}_s|} > \epsilon \right\} \right) < \epsilon \mu_{\mathcal{T}}(A) .$$

Theorem 2.20. (*Ergodic theorem for castles.*) Let $(S, \mathcal{B}, \mu, \mathcal{R}, \alpha)$ be a cross-section. Let $h : S \rightarrow \mathbb{C}$ be a bounded function. For every $\delta > 0$ there is an $\eta > 0$ and a compact set $K \subset G$ such that if \mathcal{T} is a (K, η) -invariant castle with base A such that $\mu(\text{rg}(\mathcal{T})) > \delta$, then there is a subset $B \subset A$ such that $\mu_{\mathcal{T}}(B) > (1 - \delta)\mu_{\mathcal{T}}(A)$ and such that for any $s \in B$,

$$\left| \frac{1}{|\mathcal{T}_s|} \sum_{t \in \mathcal{T}_s} h(t) - \int_S h d\mu \right| < \delta .$$

Proof. Assume that h is bounded by M and that it has zero mean. Given $\delta > 0$, let N and ϵ be as in Proposition 2.12. Given a sequence of (Følner) sets F_1, \dots, F_N , a castle \mathcal{T} , and a point s in the base of \mathcal{T} , let B_s be the set of points $t \in \mathcal{T}_s$ such that for all i , both $t \in \text{int}_{F_i} \mathcal{T}_s$ and

$$\left| \frac{1}{|\{u \mid \alpha(t, u) \in F_i\}|} \sum_{u: \alpha(t, u) \in F_i} h(u) \right| < \delta . \quad (4)$$

We call s good if $(\alpha(s, \mathcal{T}_s), \alpha(s, B_s), F_i)$ satisfies the conditions of the tiling lemma. Suppose s is good. Then we can find $t_1, \dots, t_K \in \mathcal{T}_s$ and a function $n : \{1, \dots, K\} \rightarrow \{1, \dots, N\}$ such that the sets $C_i = \{u \in \mathcal{T}_s \mid \alpha(s, u) \in F_{n(i)} \alpha(s, t_i)\}$ are δ -disjoint and cover $1 - \delta$ of \mathcal{T}_s . Let $D_i = C_i \setminus \cup_{j < i} C_j$ and let $D = \cup D_i = \cup C_i$. Then the D_i are disjoint, $|D_i| > (1 - \delta)|C_i|$, and $|D|/|\mathcal{T}_s| > 1 - \delta$. In this case,

$$\begin{aligned} \left| \frac{1}{|\mathcal{T}_s|} \sum_{t \in \mathcal{T}_s} h(t) \right| &\leq \frac{1}{|\mathcal{T}_s|} \sum_{t \in \mathcal{T}_s \setminus D} |h(t)| + \frac{1}{|\mathcal{T}_s|} \sum_i \left| \sum_{t \in D_i} h(t) \right| \leq \\ M\delta + \frac{1}{|\mathcal{T}_s|} \sum_i \left(\left| \sum_{t \in C_i} h(t) \right| + \sum_{t \in C_i \setminus D_i} |h(t)| \right) &< (2M + 1)\delta . \end{aligned}$$

We now show that we can choose Følner sequence F_i , a compact set $K \subset G$, and $\eta > 0$ such that if \mathcal{T} is a (K, η) -invariant castle, then

most (in the sense of $\mu_{\mathcal{T}}$ -measure) of the points of the base of \mathcal{T} are good. First note that for every compact set U , if F is $(U, 1)$ -invariant and $X \subset F$ is the set of centers of disjoint right translates of U , then

$$|X|\lambda(U) = \lambda(XU) \leq \lambda(FU) \leq 2\lambda(F),$$

and so $|X| \leq \frac{2}{\lambda(U)}\lambda(F)$.

It is therefore enough to show that for most points s in the base of \mathcal{T} and for most points $t \in \mathcal{T}_s$,

1. $\{u|\alpha(t, u) \in F_i\} \subset \mathcal{T}_s$ for all i .
2. (4) holds.
3. $|\{u \in \mathcal{T}_s | \alpha(t, u) \in (\cup_{j < i} F_j)^{-1} F_i\}| < (1 + \epsilon)|\{u \in \mathcal{T}_s | \alpha(t, u) \in F_i\}|$.

Choose the F_i to be sufficiently invariant so that the mean ergodic theorem is satisfied for $1 - \delta\epsilon^2$ of the points in S , for each F_i , and for the functions h and 1. By Fubini, 2. is satisfied for at least $1 - \epsilon^2$ of the points in $rg(\mathcal{T})$. By Fubini again, for $1 - \epsilon$ of the points of the base, 2. is satisfied for $1 - \epsilon$ of the points above them. We require additionally that F_i is $((\cup_{j < i} F_j)^{-1}, \frac{\lambda(U)}{2}\epsilon)$ -invariant, and denote $L_i = (\cup_{j < i} F_j)^{-1} F_i \setminus F_i$. We can further require that L_i is $(U, 1)$ -invariant, and so for all $t \in \mathcal{T}_s$, $|\{u \in \mathcal{T}_s | \alpha(t, u) \in L_i\}| < c\lambda(L_i)$, where $c = \frac{2}{\lambda(U)}$. By the ergodic theorem for the function 1, $|\{u | \alpha(t, u) \in F_i\}| > (1 - \epsilon)\lambda(F_i)$. Now we add the demand that \mathcal{T} is $(\cup_i F_i, \epsilon)$ -invariant. Then 1. is satisfied for most pairs (s, t) , and for each point $t \in \text{int}_{F_i} \mathcal{T}_s$,

$$\frac{|\{u \in \mathcal{T}_s | \alpha(t, u) \in L_i\}|}{|\{u \in \mathcal{T}_s | \alpha(t, u) \in F_i\}|} < \frac{c\lambda(L_i)}{(1 - \epsilon)\lambda(F_i)} < \epsilon$$

which implies 3. □

3 Entropy Theory

3.1 Entropy Theory for Amenable Groups Actions

The entropy theory of actions of amenable groups is developed in [OW]. The definition of entropy takes simple form for groups with zero self entropy, which are defined as follows:

Definition 3.1. Let G be a unimodular amenable group, and choose a right invariant metric, d , on G that generates the topology. We say that G has zero self entropy if the following condition holds: For every compact subset K of G , and for every $\epsilon > 0$, if $F \subset G$ is sufficiently invariant, then there is a partition of K into less than $2^{\epsilon\lambda(F)}$ sets such that for every two points g_1, g_2 in the same part and every $f \in F$, we have $d(fg_1, fg_2) < \epsilon$.

Example 3.2. 1. Every nilpotent Lie group has zero self entropy.
2. There are unimodular solvable Lie groups that do not have zero self entropy.

We now briefly remind the reader the definition of entropy for actions of groups with zero self entropy. We are given an action of the group G on a probability space X . A partition of X is a function P from X to a finite set, whose size we denote by $|P|$. For every partition P of X and a compact set $F \subset G$, we say that $x, y \in X$ are (P, F, ϵ) -close if

$$\lambda(\{f \in F | P(fx) \neq P(fy)\}) < \epsilon\lambda(F).$$

A (P, F, ϵ) -ball is a subset $A \subset X$ such that every two points in A are (P, F, ϵ) -close. The entropy of the action with respect to P is defined to be the minimal number $h = h(G, P)$ such that for every $\epsilon > 0$, if $F \subset G$ is sufficiently invariant, then there is a collection \mathcal{C} of (P, F, ϵ) -balls that covers $1 - \epsilon$ of X and has size less than $2^{(h+\epsilon)\lambda(F)}$. We define the entropy of the action of G on X , denoted by $h(X)$, to be the supremum of the numbers $h(X, P)$, where P is taken from the set of partitions of X .

Remark 3.3. 1. If X is a Borel space endowed with regular probability measure, then $h(X)$ is the supremum of the numbers $h(X, Q)$, where Q is taken from the set of partitions of X such that any part of Q has negligible boundary. This is true since this collection of partitions generates the sigma-algebra.
2. In the other direction, one can consider more general partitions, which are (Borel) maps to some compact metric space (Z, d) . If P is such a generalized partition, we say that two points $x, y \in X$ are (P, F, ϵ) -close if

$$\int_F d(P(fx), P(fy)) < \epsilon\lambda(F),$$

and define similarly the notions of (P, F, ϵ) -ball and the entropy of the action with respect to P . It turns out (see [OW]) that the supremum of $h(X, P)$ where P is taken from the set of generalized partitions is again $h(X)$.

3.2 Entropy Theory for Cross Sections

In our definition of entropy for cross-sections, we choose a slightly different route, concentrating on castles.

Definition 3.4. Let $(S, \mathcal{B}, \mu, \mathcal{R}, \alpha)$ be a cross-section and let \mathcal{G} be a sub-sigma-algebra of \mathcal{B} . Let P be a finite partition of S , not necessarily \mathcal{G} -measurable, let \mathcal{T} be a castle with base A , and let $\epsilon > 0$. A $(P, \mathcal{T}, \epsilon, \mathcal{G})$ -ball is a triple (B, E, ϕ) such that B is a subset of A , not necessarily \mathcal{G} -measurable, E is a finite set, and $\phi : B \times E \rightarrow S$ is a restriction of a \mathcal{G} -measurable, one-to-one function such that

1. $\phi(x, e) \in \mathcal{T}_x$ for all $x \in B$, $e \in E$.
2. $|\phi(x, E)| > (1 - \epsilon)|\mathcal{T}_x|$ for every $x \in B$.
3. $x \mapsto P(\phi(x, e))$ is constant for every $e \in E$.

Given (B, E, ϕ) as above, for any $s, t \in B$ we get a partially defined map $\phi_{s,t}$ between \mathcal{T}_s and \mathcal{T}_t by $\phi_{s,t}(r) = \phi(t, \pi_2(\phi^{-1}(r)))$ where π_2 is the projection to the second coordinate. The function $\phi_{s,t}$ is defined for at least $1 - \epsilon$ of the points of \mathcal{T}_s , and its image consists of at least $1 - \epsilon$ of the points of \mathcal{T}_t . Conversely, if B is given and we can show that for every $s, t \in B$ there is such a $\phi_{s,t}$ that depends on s and t in a \mathcal{G} -measurable way, then B can be decomposed in a \mathcal{G} -measurable way into balls—that is, there are \mathcal{G} -measurable sets C_1, \dots, C_N , such that $B \subset \cup C_i$, and there are E_i, ϕ_i as above such that $(B \cap C_i, E_i, \phi_i)$ are $(P, \mathcal{T}, 2\epsilon, P)$ -balls.

Definition 3.5. Let \mathcal{G} be a sigma-algebra, let $A \subset S$ be a \mathcal{G} -measurable set, and let $\mathcal{C} = \{C_1, \dots, C_N\}$ be a collection of Borel sets (not necessarily \mathcal{G} -measurable) contained in A . Denote $\log_+(x) = \max\{\log(x), 0\}$. The relative logarithmic size of \mathcal{C} with respect to \mathcal{G} is

$$rls_A(\mathcal{C}|\mathcal{G}) = \int_A \log_+(|\{i | E(1_{C_i}|\mathcal{G})(x) > 0\}|) d\mu(x).$$

Remark 3.6. One can change the sets C_i by sets of measure 0 such that the integrand $|\{i | E(1_{C_i}|\mathcal{G})(x) > 0\}|$ is the number of sets in \mathcal{C} that intersect the \mathcal{G} -fiber of x .

Definition 3.7. Suppose that \mathcal{T} is a \mathcal{G} -measurable castle with base A . Let P be a partition. Denote by $h_{\mathcal{T}}^{\epsilon}(P|\mathcal{G})$ the infimum of $rls_A(\mathcal{C})$ where \mathcal{C} is a collection of $(P, \mathcal{T}, \epsilon, \mathcal{G})$ -balls that cover A .

Remark 3.8. Clearly, $h_{\mathcal{T}}^{\epsilon}$ is monotone in \mathcal{T} —if \mathcal{T} and \mathcal{T}' are \mathcal{G} -measurable castles with disjoint ranges, then for every partition P and $\epsilon > 0$, we have $h_{\mathcal{T} \sqcup \mathcal{T}'}^{\epsilon}(P|\mathcal{G}) = h_{\mathcal{T}}^{\epsilon}(P|\mathcal{G}) + h_{\mathcal{T}'}^{\epsilon}(P|\mathcal{G})$. Also, if $|P|$ is a set whose size is less than N , then $h_{\mathcal{T}}^{\epsilon}(P|\mathcal{G}) < \mu(\text{rg}(\mathcal{T})) \log N$.

Theorem 3.9. Let $(S, \mathcal{B}, \mu, \mathcal{R}, \alpha)$ be a cross-section. Suppose that \mathcal{G} is a sigma-algebra such that α is \mathcal{G} -measurable. Let $0 < \epsilon < 1$. Suppose that \mathcal{T} is a tower that covers more than $1 - \epsilon$ of S . There is a compact set $L \subset G$ and a $\delta > 0$ such that if \mathcal{T}' is an (L, δ) -invariant castle that is \mathcal{G} -measurable and covers $1 - \epsilon$ of S , then $h_{\mathcal{T}'}^{4\epsilon}(P|\mathcal{G}) \leq h_{\mathcal{T}}^{\epsilon}(P|\mathcal{G}) + 2\epsilon + \epsilon \log(|P|)$.

Proof. Let A be the base of \mathcal{T} . There is a compact subset $K \subset G$ such that for all s in A , except for a subset of $\mu_{\mathcal{T}}$ -measure as small as we want, $\alpha(s, \mathcal{T}_s) \subset K$. After removing this small set (and noting that we can choose K to be so large so that we still have a castle that covers more than $1 - \epsilon$ of S), we may assume that $\alpha(s, \mathcal{T}_s) \subset K$ for all s in A . Let \mathcal{C} be a collection of $(P, \mathcal{T}, \epsilon, \mathcal{G})$ -balls that covers A , such that $rls_A(\mathcal{C}|\mathcal{G}) < h_{\mathcal{T}}^{\epsilon}(P|\mathcal{G}) + \epsilon$. Let $H(x) = \log_+(|\{i | E(1_{C_i}|\mathcal{G})(x) > 0\}|)$ (so $H(x) = 0$ outside A). By the ergodic theorem for castles, if \mathcal{T}' is a sufficiently invariant castle with base A' that covers more than ϵ of S , then for a subset $B \subset A'$ of $\mu_{\mathcal{T}'}$ -measure greater than $1 - \epsilon$, we have that for every $s \in B$,

$$\left| \frac{1}{|\mathcal{T}'_s|} \sum_{t \in \mathcal{T}'_s} H(t) - \int_S H d\mu \right| < \epsilon \quad (5)$$

and

$$\left| \frac{1}{|\mathcal{T}'_s|} \sum_{t \in \mathcal{T}'_s} 1_{\text{rg}(\mathcal{T})}(t) - \mu(\text{rg}(\mathcal{T})) \right| < \epsilon. \quad (6)$$

Taking \mathcal{T}' as sufficiently invariant, and by making B smaller, we can assume that for all points s of B , both (5) and (6) hold, that \mathcal{T}'_s is (K, ϵ) -invariant, and that $\mu_{\mathcal{T}'}(B) > (1 - \epsilon)\mu_{\mathcal{T}'}(A)$.

Since \mathcal{T}' and B are \mathcal{G} -measurable, it is possible to divide B into \mathcal{G} -measurable subsets, B_i , and to find a one-to-one and measure-preserving map $\phi_i : B_i \times E_i \rightarrow \mathcal{T}'$, where E_i are finite sets, such

that $\phi_i(s, E_i) = \mathcal{T}'_s$ and the following holds: if we denote for every $s, t \in B_i$ the identification $t \mapsto \phi_i(s_2, \pi_2(\phi_i^{-1}(t)))$ between \mathcal{T}'_s and \mathcal{T}'_t by $\phi_{s,t}$, then $E(1_{C_i}|\mathcal{G})(r) > 0$ iff $E(1_{C_i}|\mathcal{G})(\phi_{s,t}(r)) > 0$ for all $s, t \in B_i$ and $r \in \mathcal{T}'_s$. Let $s \sim t$ if, for every $r \in \mathcal{T}'_s$, the points r and $\phi_{s,t}(r)$ are contained in the same C_i . Denote the equivalence classes inside B_j by $\mathcal{D}_1^j, \dots, \mathcal{D}_{N_j}^j$. If $s \in B_j$, the number N_j of equivalence classes is $\exp(\sum_{t \in \mathcal{T}_s} H(t))$, and by (5), this number is less than $\exp((h_{\mathcal{T}}(P|\mathcal{G}) + 2\epsilon)|\mathcal{T}'_s|)$. Therefore, the relative logarithmic size of the collection $\{\mathcal{D}_i^j\}$ over B is less than $h_{\mathcal{T}}(P|\mathcal{G}) + 2\epsilon$.

Next, we show that each \mathcal{D}_i^j is a $(P, \mathcal{T}', 4\epsilon, \mathcal{G})$ -ball. For $s \in \mathcal{D}_i^j$, denote $R_s = \{r \in \mathcal{T}'_s | r \in A, \mathcal{T}_r \subset \mathcal{T}'_s\}$. From (6) and the assumption that \mathcal{T}'_s is (K, ϵ) -invariant, we get that for all $s \in B$,

$$\frac{|\bigcup_{r \in R_s} \mathcal{T}_r|}{|\mathcal{T}'_s|} > 1 - 3\epsilon.$$

Given s_1, s_2 in B , we define a map $\psi : \mathcal{T}'_{s_1} \rightarrow \mathcal{T}'_{s_2}$. On R_{s_1} it is equal to ϕ_{s_1, s_2} . Given $r \in R_{s_1}$, since r and $\phi_{s_1, s_2}(r)$ are in the same C_i , there exists a function $\theta_r : \mathcal{T}_r \rightarrow \mathcal{T}_{\phi_{s_1, s_2}(r)}$ such that $P(t) = P(\theta_r(t))$ for $1 - \epsilon$ of the points in \mathcal{T}_r . Since the sets \mathcal{T}_r for $r \in R_s$ are disjoint, we can define a function $\theta : \bigcup_{r \in R_{s_1}} \mathcal{T}_r \rightarrow \mathcal{T}'_{s_2}$ such that $P(\theta(t)) = P(t)$ for $1 - \epsilon$ of the points $t \in \bigcup_{r \in R_{s_1}} \mathcal{T}_r$. Extending this function in an arbitrary way to the rest of \mathcal{T}'_{s_1} , we get a \mathcal{G} -measurable function such that $P(w) = P(\theta(w))$ for $1 - 4\epsilon$ of the points.

Therefore, if we denote the restriction of \mathcal{T}' to B by $\mathcal{T}'|_B$, we get that

$$h_{\mathcal{T}'}^{4\epsilon}(P|\mathcal{G}) \leq h_{\mathcal{T}'|_B}^{4\epsilon}(P|\mathcal{G}) + \mu(\text{rg}(\mathcal{T}'_{A' \setminus B})) \log |P| \leq h_{\mathcal{T}}(P|\mathcal{G}) + 2\epsilon + \epsilon \log |P|.$$

□

Definition 3.10. Let S be a cross-section, let P be a partition and let \mathcal{G} be a sub-sigma-algebra. We define the relative entropy to be

$$h(P|\mathcal{G}) = \lim_{\epsilon \rightarrow 0} \lim_{\mathcal{T}} h_{\mathcal{T}}^{\epsilon}(P|\mathcal{G})$$

where the inner limit is taken over more and more invariant castles that cover $1 - \epsilon$ of S .

Remark 3.11. By the previous theorem, the limit above exists.

The following is an analogue of Shannon–McMillan’s theorem.

Theorem 3.12. *Suppose $(S, \mathcal{B}, \mu, \mathcal{R}, \alpha)$ is a U -discrete cross-section. Let P be a partition of S , and let $\mathcal{G} \subset \mathcal{B}$ be a sub-sigma-algebra such that α is \mathcal{G} -measurable. Then $h(P|\mathcal{G})$ is the infimum over the positive real numbers h , such that for every $\epsilon > 0$, for sufficiently invariant Følner sets $F \subset G$, there is a collection $\mathcal{C} = \{C_i, \dots, C_N\}$ of subsets of S , finite sets E_1, \dots, E_N , and one-to-one functions $\phi_i : C_i \times E_i \rightarrow S$, that are restrictions of \mathcal{G} -measurable functions, such that*

1. *For each i and $s \in C_i$, we have that $\phi_i(s, E_i) \subset \{t | \alpha(s, t) \in F\}$ and $|\phi(s, E_i)| > (1 - \epsilon)|\{t | \alpha(s, t) \in F\}|$.*
2. *For each i and $e \in E_i$, the function $s \mapsto P(\phi_i(s, e))$ ($s \in C_i$) is constant.*
3. *$|\{i | E(1_{C_i}|\mathcal{G})(s) > 0\}| < \exp((h + \epsilon)|\{t | \alpha(s, t) \in F\}|)$ for all $s \in S$.*
4. *$\mu(\bigcup_i C_i) > 1 - \epsilon$.*

Proof. Choose a neighborhood V such that $V^2 \subset U$. In one direction, assume that h is such that for any $\epsilon > 0$, if F is sufficiently invariant, then there are C_i, E_i, ϕ_i as above. We show that for any $\epsilon > 0$, if \mathcal{T} is a sufficiently invariant castle that covers $1 - \epsilon$ of S , then $h_{\mathcal{T}}^{\epsilon}(P|\mathcal{G}) < h + \epsilon$.

By Proposition 2.12, there is $\eta > 0$ and $n > 0$ such that if $A, B, F_1, \dots, F_n \subset G$ satisfy the conditions of the tiling lemma with parameter η , then the conclusions of Proposition 2.12 hold with δ replaced by ϵ . Choose F_1, \dots, F_n such that

- F_i is $(U, 1)$ -invariant.
- $\lambda(F_i) \geq 0$.
- For every i , $(\bigcup_{j < i} F_j)^{-1} F_i \setminus F_i$ is $(U, 1)$ -invariant.
- For every i , $\lambda((\bigcup_{j < i} F_j)^{-1} F_i \setminus F_i) < \eta \lambda(F_i)$.

For every $1 \leq k \leq n$, let $\mathcal{C}^k = \{C_1^k, \dots, C_{m(k)}^k\}$ be a collection of subsets of S , let E_i^k , $1 \leq i \leq m(k)$, be a collection of finite sets, and let $\phi_i^k : C_i^k \times E_i^k \rightarrow S$, $1 \leq i \leq m(k)$, be a collection of one-to-one functions that are the restrictions of \mathcal{G} -measurable functions, such that 1., 2., 3., and 4. in the conditions of the theorem are satisfied with ϵ replaced by $\epsilon \delta / n$.

By the ergodic theorem, if \mathcal{T} is sufficiently invariant, then for $1 - \epsilon$ of the points s in the base of \mathcal{T} we have that at least a $1 - \eta$ portion of \mathcal{T}_s is contained in $\bigcap_{k=1}^n \bigcup \mathcal{C}^k$. For each such a point, apply Proposition 2.12 to find

- A number $K(s)$.
- Points $x_i(s) \in \mathcal{T}_s \cap (\cap_k \cup \mathcal{C}^k)$, for $1 \leq i \leq K(s)$.
- Indices $j(i, s) \in \{1, \dots, n\}$.

such that the sets $\{y | \alpha(x_i(s), y) \in F_{j(i, s)}\}$ are ϵ -disjoint and cover $1 - \epsilon$ of \mathcal{T}_s . From the proof of Proposition 2.12, we can assume that the functions $K(-)$, $x_i(-)$, and $j(i, -)$ are \mathcal{G} -measurable. For two points, s and t , in the base of \mathcal{T} for which the above holds, we say that s is equivalent to t if $K(s) = K(t)$, for all $1 \leq i \leq K(s)$, $j(i, s) = j(i, t)$, and for all $1 \leq i \leq K(s)$, the points $x_i(s)$ and $x_i(t)$ belong to the same set in $\mathcal{C}^{j(i, s)}$. It is easy to see that each equivalence class is a $(P, \mathcal{T}, 2\epsilon, \mathcal{G})$ -ball. Moreover, for any s , the number of equivalence classes that intersect the \mathcal{G} -fiber of s is at most

$$\begin{aligned} & \prod_{i=1}^{K(s)} \left| \{k | E(1_{\mathcal{C}_k^{j(i, s)}} | \mathcal{G})(x_i(s)) > 0\} \right| \leq \\ & \leq \exp \left((h + \epsilon) \sum_{i=1}^{K(s)} |\{t | \alpha(x_i(s), t) \in F_{j(i, s)}\}| \right) \leq \\ & \leq \exp((h + \epsilon)(|\mathcal{T}_s| + \epsilon)) \leq \exp((h + 2\epsilon)|\mathcal{T}_s|), \end{aligned}$$

and therefore $h_{\mathcal{T}}^{2\epsilon}(P|\mathcal{G}) \leq h + 2\epsilon$.

In the other direction, let \mathcal{T} be a \mathcal{G} -measurable castle with base A that covers $1 - \epsilon$ of S . After discarding a small enough subset of the base of \mathcal{T} , we can assume that the set $\{\alpha(s, t) | (s, t) \in \mathcal{T}\}$ is contained in some compact set L . Suppose that there is a collection $\mathcal{D} = \{D_1, \dots, D_n\}$ of $(P, \mathcal{T}, \epsilon, \mathcal{G})$ -balls that covers $1 - \epsilon$ of A , such that $rls_A \mathcal{D} < h$. Let $H : S \rightarrow \mathbb{R}$ be the function such that $H(s)$ is equal to $\log_+ (|\{i | E(1_{D_i} | \mathcal{G})(s) > 0\}|)$ if s belongs to A , and is equal to 0 if s is not in A . We have that $\int_S H d\mu = rls_A \mathcal{D} < h$. Let $F \subset G$ be sufficiently invariant, such that for $1 - \epsilon$ of the points s of S ,

- $(1 - \epsilon)\lambda(F) \leq |\{t | \alpha(s, t) \in F\}| \leq (1 + \epsilon)\lambda(F)$.
- $|\{t \in S | \alpha(s, t) \in F \wedge t \in rg(\mathcal{T})\}| > (1 - \epsilon)|\{t \in S | \alpha(s, t) \in F\}|$.
- $\sum_{t: \alpha(s, t) \in F} H(t) \leq (h + \epsilon)|\{t | \alpha(s, t) \in F\}|$.

The existence of such F is given by the ergodic theorem.

By enlarging F , we can assume that F is (L, ϵ) -invariant. For all but $1 - 2\epsilon$ of the points s of S , there is a number $K(s)$ and elements $x_i(s) \in S$, $1 \leq i \leq K(s)$ such that

- For every i , the point $x_i(s)$ belongs to the base of \mathcal{T} .
- The sets $\mathcal{T}_{x_i(s)}$ cover $1 - 2\epsilon$ of $\{t | \alpha(s, t) \in F\}$.
- $\sum_{i=1}^{K(s)} H(x_i(s)) \leq (h + 2\epsilon) |\{t | \alpha(s, t) \in F\}|$.

We can assume, moreover, that the functions $K(-)$ and $x_i(-)$ are \mathcal{G} -measurable. Define an equivalence relation on the set of points of S for which the above holds: $s \sim t$ if $K(s) = K(t)$ and for all $1 \leq i \leq K(s)$, the points $x_i(s)$ and $x_i(t)$ belong to the same set in \mathcal{D} .

Let $\mathcal{C} = \{C_i\}$ be the collection of equivalence classes. For every K , the number of equivalence classes that are contained in $\{s | K(s) = K\}$ is less than $2^{(h+3\epsilon)\lambda(F)}$. Therefore, for every $s \in S$,

$$|\{i | E(1_{C_i} | \mathcal{G})(s) > 0\}| \leq 2^{(h+3\epsilon)\lambda(F)} \leq 2^{(h+4\epsilon)|\{t | \alpha(s, t) \in F\}|}.$$

For every two equivalent points s, t , there is a map $\phi_{s,t} : \{u | \alpha(s, u) \in F\} \rightarrow \{t | \alpha(t, v) \in F\}$, such that, for $1 - 3\epsilon$ of the points in $\{u | \alpha(s, u) \in F\}$, we have $P(\phi_{s,t}(u)) = P(u)$. As explained after Definition 3.4, there is a \mathcal{G} -measurable refinement of the C_i 's, such that, for each new part Y , there is a set E and a function $\phi : X \times E \rightarrow S$, which is the restriction of a \mathcal{G} -measurable function, such that 1. and 2. in the theorem hold. Since the refinement is \mathcal{G} -measurable, the relative size of it is unchanged, and so 3. holds. Lastly, 4. holds because the union is just the union of \mathcal{C} .

□

Theorem 3.13. *Let \mathcal{R} be a probability-preserving equivalence relation on the probability space (S, \mathcal{B}, μ) . Let \mathcal{G} be a sub-sigma-algebra of \mathcal{B} , and suppose that there are two amenable groups G, H , with zero self entropy, and two \mathcal{G} -measurable cocycles $\alpha : \mathcal{R} \rightarrow G$ and $\beta : \mathcal{R} \rightarrow H$ such that $(S, \mathcal{B}, \mu, \mathcal{R}, \alpha)$ and $(S, \mathcal{B}, \mu, \mathcal{R}, \beta)$ satisfy the mean ergodic theorem. For every partition P , denote by $h_\alpha(P | \mathcal{G})$ the relative entropy of P with respect to \mathcal{G} for the cross-section $(S, \mathcal{B}, \mu, \mathcal{R}, \alpha)$, and define similarly $h_\beta(P | \mathcal{G})$. Then for each partition P , $h_\alpha(P | \mathcal{G}) = h_\beta(P | \mathcal{G})$.*

Proof. Let $S, \mathcal{B}, \mu, \mathcal{R}, \alpha, \beta, \mathcal{G}$ be as in the statement of the theorem.

Lemma 3.14. *For every $\epsilon > 0$ and a compact set $L \subset H$, there is a compact set $F \subset G$ such that if \mathcal{T} is a castle, and is (F, ϵ) -invariant, then it is also $(L, 2\epsilon)$ -invariant.*

Proof. Let $K \subset G$ be a compact set such that the set

$$Z = \{x | \text{for every } y \text{ such that } \beta(x, y) \in L, \text{ we have } \alpha(x, y) \in F\}$$

has measure greater than $1 - \epsilon/2$. By the ergodic theorem, there is a compact set $F \subset G$ that contains K such that if \mathcal{T} is (F, ϵ) -invariant, then $1 - \epsilon$ of the points in the range of \mathcal{T} are in Z . But if s is in the base of \mathcal{T} and $x \in \mathcal{T}_s$ is both in Z and in the F -interior of \mathcal{T}_s , then x is also in the L -interior of \mathcal{T}_s . \square

Let $L_n \subset H$ be a Følner sequence in H . Take a Følner sequence $F_n \subset G$ that satisfies the conclusion of the lemma.

The definition of a $(P, \mathcal{T}, \epsilon, \mathcal{G})$ -ball does not depend on the cocycle, and hence neither is the definition of $h_{\mathcal{T}}^{\epsilon}(P|\mathcal{G})$. The number $h_{\alpha}(P|\mathcal{G})$ is the limit of $h_{\mathcal{T}}^{\epsilon}(P|\mathcal{G})$ as \mathcal{T} is taken to be (L_n, ϵ) -invariant, and $h_{\beta}(P|\mathcal{G})$ is the limit of $h_{\mathcal{T}}^{2\epsilon}(P|\mathcal{G})$ as \mathcal{T} is taken to be $(F_n, 2\epsilon)$ -invariant. These limits are equal. \square

The following is a version of Abramov's theorem.

Theorem 3.15. *Let G be an amenable group of zero self entropy. Assume G acts on a Borel probability space (X, \mathcal{B}, m) , and that $S \subset X$ is a U -discrete cross-section, for some neighborhood U of 1 in G . Let \mathcal{G} is a G -invariant sub-sigma-algebra such that S is \mathcal{G} -measurable. Then*

$$h(G, X|\mathcal{G}) = h(S|\mathcal{G}_S)$$

where \mathcal{G}_S denotes the restriction of \mathcal{G} to S .

- Remark 3.16.** 1. On the left hand side, the relative entropy $h(G, X|\mathcal{G})$ is defined as the difference of the entropy of the action of G on (X, \mathcal{B}) and the entropy of the action of G on (X, \mathcal{G}) . Since G has zero self entropy, the relative entropy is the difference of the spatial entropies (see [OW, II.4]). For our argument, we need that the relative entropy can be approximated by choosing a large Følner set L , and a \mathcal{G} -measurable partition of X , such that most parts can be partitioned into $2^{(1+\epsilon)\lambda(L)h(G, P|\mathcal{G})}$ (P, L, ϵ) -balls.
2. The formula is simpler than in the classical Abramov theorem because of our normalization 2.4 of the Haar measure.

Proof. Let P be a partition of X , and let $\epsilon > 0$. We assume that every part of P has negligible boundary. Choose a compact set $F \subset G$ such that $m(F \cdot S) > 1 - \epsilon/2$. Let $\eta = \frac{\epsilon}{2\lambda(F)}$. Define a generalized partition \tilde{P} as follows: the value set of the partition is the set of functions from F to $|P|$, with metric given by

$$d(\phi, \psi) = \lambda\{f \in F | \phi(f) \neq \psi(f)\},$$

and $\tilde{P}(x)$ is the function $f \mapsto P(fx)$. Let Ω be the image of \tilde{P} . Since \tilde{P} is continuous and X is compact, Ω is also compact. Choose an $\eta/2$ -net \mathcal{A} in Ω . Choose also a measurable map $\rho : \Omega \rightarrow \mathcal{A}$ such that $d(\omega, \rho(\omega)) < \eta/2$ for all $\omega \in \Omega$. Let $\tilde{\tilde{P}}$ be the partition of S given by $\tilde{\tilde{P}}(s) = \rho(\tilde{P}(s))$.

Let W be a neighborhood of the identity in G such that the set

$$B = \{x \in X \mid \text{the function } w \mapsto P(wx) \text{ is constant on } w \in W\}$$

has measure greater than $1 - \epsilon/2$.

Finally, choose $L \subset G$ compact such that L is (F, ϵ) -invariant, and for all $s \in S$, outside a set of μ -measure less than ϵ , the following hold:

1. $\lambda\{f \in L \mid fs \in B\} > (1 - \epsilon)\lambda(L)$.
2. $\lambda\{f \in L \mid fs \notin F \cdot S\} < \epsilon\lambda(L)$.
3. $|\{f \in L \mid fs \in S\}| < 2\lambda(L)$.
4. (Theorem 3.12) There is a collection $\mathcal{C} = \{C_i\}$ of subsets of S such that $rls(\mathcal{C}|\mathcal{G}) < \lambda(L)(h(S|\mathcal{G}) + \epsilon)$, such that $\mu(\cup \mathcal{C}) > 1 - \epsilon$, and such that each C_i is a $(\tilde{\tilde{P}}, L, \eta)$ -ball.
5. (Zero self-entropy of G) There is a partition of F into less than $2^{\epsilon\lambda(L)}$ parts M_i such that for g_1, g_2 in the same M_i and for any $f \in L$, there is $\delta(f, g_1, g_2) \in W$ such that $fg_1 = \delta(f, g_1, g_2)fg_2$.

Lemma 3.17. *There is a \mathcal{G} -measurable partition \mathcal{D} such that every set in the common refinement $\mathcal{C} \vee \mathcal{D}$ is a $(P, L, 3\epsilon)$ -ball.*

Proof. Suppose first that $x_1, x_2 \in C_i$ have the same return times to S , i.e.

$$A := \{f \in L \mid fx_1 \in S\} = \{f \in L \mid fx_2 \in S\}.$$

By assumption, there is a set $B \subset A$ such that $|B| > (1 - \eta)|A|$ and for every $f \in B$, $\tilde{\tilde{P}}(fx_1) = \tilde{\tilde{P}}(fx_2)$. If we denote by 1_ϕ the characteristic function of ϕ , then

$$\begin{aligned} \lambda\{f \in L \mid P(fx_1) \neq P(fx_2)\} &= \int_L 1_{P(fx_1) \neq P(fx_2)} df \leq \\ &\lambda(L \setminus FA) + \lambda(FA \setminus FB) + \sum_{b \in B} \int_F 1_{P(fb_{x_1}) \neq P(fb_{x_2})} df. \end{aligned}$$

The first summand is less than $\epsilon\lambda(L)$ by assumption 2 on L . The second summand is less than

$$|A \setminus B|\lambda(F) < \eta|A|\lambda(F) \leq \epsilon\lambda(L),$$

and since for every $b \in B$, the distance between the (\tilde{P}, F) -names of bx_1 and bx_2 is less than η , we get that the third summand is less than $|B|\eta \leq \epsilon\lambda(L)/2$. Therefore,

$$\lambda\{f \in L | P(fx_1) \neq P(fx_2)\} < 2.5\epsilon\lambda(L).$$

Since the cocycle α is \mathcal{G} -measurable, for every small enough neighborhood $Z \subset G$ of the identity, there is a \mathcal{G} -measurable partition \mathcal{D}_Z such that every x_1, x_2 in the same part and every $f \in L$ such that $fx_1 \in S$, there is a unique $g \in Z$ such that $gfx_2 \in S$. By continuity, there is a neighborhood Z such that every x_1, x_2 in the same atom of $\mathcal{C} \vee \mathcal{D}_Z$,

$$\lambda\{f \in L | P(fx_1) \neq P(fx_2)\} < 3\epsilon\lambda(L).$$

□

Since $rls(\mathcal{C} \vee \mathcal{D} | \mathcal{G}) = rls(\mathcal{C} | \mathcal{G})$, we can replace \mathcal{C} by $\mathcal{C} \vee \mathcal{D}$ and assume that each C_i is a $(P, L, 3\epsilon)$ -ball. We show now that each $M_i C_j$ is a $(P, L, 5\epsilon)$ -ball. Indeed, if $x_1, x_2 \in C_i$ and $g_1, g_2 \in M_i$, then the distance between the (L, P) -names of $g_1 x_1$ and $g_2 x_2$ is

$$\begin{aligned} \lambda\{f \in L | P(fg_1 x_1) \neq P(fg_2 x_2)\} &\leq \lambda\{f \in L | P(fg_1 x_1) \neq P(fg_2 x_1)\} + \\ &+ \lambda\{f \in L | P(fg_2 x_1) \neq P(fg_2 x_2)\}. \end{aligned}$$

By the definition of M_i , we have that $fg_1 = \delta(f, g_1, g_2)fg_2$, where $\delta(f, g_1, g_2) \in W$. Hence, the first summand is less than $\lambda\{f \in L | fg_2 x_1 \notin B\} < \epsilon\lambda(L)$. Since L is (F, ϵ) -invariant, the second summand is less than $\epsilon\lambda(L) + \lambda\{f \in L | P(fx_1) \neq P(fx_2)\} < 4\epsilon\lambda(L)$. We get that $h(G, X | \mathcal{G}) \leq h(S | \mathcal{G}_S)$.

For the other direction, we start with a definition.

Definition 3.18. Given a partition P of S and $U \subset G$ open neighborhood of 1, the U -fattening of P is the partition Q of X into $|P| + 1$ parts defined as follows: $Q(us) = P(s)$ if $u \in U, s \in S$, and $X \setminus US$ is the last part.

Suppose we have a partition \tilde{P} of S . Assume that S is U -discrete for a neighborhood U of the identity in G . Without loss of generality we can assume that $\lambda(U) < 1$. Let P be the U fattening of \tilde{P} to X . By definition, there is a Følner set F and a partition $\mathcal{C} = \{C_1, \dots, C_N\}$ of $1 - \epsilon\lambda(U)$ of X such that for every $x \in X$, the set of i 's such that $E(1_{C_i}|\mathcal{G})(x) > 0$ has size less than $\exp((h + \epsilon)\lambda(F))$, and for every two points x, y in the same part

$$\lambda(\{f \in F | P(fx) \neq P(fy)\}) < \epsilon\lambda(F).$$

We can assume that F is (U, ϵ) -invariant and get that for every $u \in U$,

$$\lambda(\{f \in F | P(fux) \neq P(fuy)\}) < 2\epsilon\lambda(F).$$

By Fubini, there is $u \in U$ such that \mathcal{C} covers more than $1 - \epsilon$ of uS . It is now easy to check that $u^{-1}C_i \cap S$ is a collection of $(P, F, 3\epsilon)$ -balls. Since we did not increase its size, Theorem 3.12 implies that $h(S|\mathcal{G}_S) \leq h(G, X|\mathcal{G}) + 3\epsilon$. □

4 CPE actions

We recall some definitions.

Definition 4.1. Let G be an amenable group with zero self entropy. An action of G on a probability space X is called completely positive entropy (or CPE for short) if, for any non-trivial partition P of X , the entropy $h(G, P)$ is positive.

Definition 4.2. An action of an amenable group G on a probability space X is called uniformly mixing if, for every partition P of X and any $\epsilon > 0$, there is a compact $K \subset G$ such that for any finite set $F \subset G$ that is K -separated (i.e. $gh^{-1} \notin K$ for any two distinct $g, h \in F$), one has

$$\left| \frac{1}{|F|} H \left(\bigvee_{g \in F} gP \right) - H(P) \right| < \epsilon. \quad (7)$$

Definition 4.3. Let G be an amenable group acting on a probability space X . The spectrum of this action is the associated G -representation on $L_2(X)$ given by

$$(\rho(g)f)(x) = f(g^{-1}x).$$

The spectrum is called Lebesgue with multiplicity N (which can be infinity) if $L_2(X)$ decomposes into direct sum of N copies of the regular representation of G .

Theorem 4.4. *Let G be an amenable group with zero self entropy, and let X be a probability space. Suppose that G acts on X in a CPE manner. Then*

1. *The action is uniformly mixing.*
2. *The spectrum of the action is Lebesgue with multiplicity \aleph_0 .*

Proof. We first relativize the notions of CPE, uniform mixing, and spectrum as follows:

Definition 4.5. Let \mathcal{G} be an invariant sub-sigma-algebra. The action of G on X is said to be relatively CPE over \mathcal{G} if, for any partition P , which is not measurable with respect to \mathcal{G} , we have $h(G, P|\mathcal{G}) > 0$. We say that the action is uniformly mixing relative to \mathcal{G} if the inequality in 4.2 holds after we replace all entropies with relative entropies (with respect to \mathcal{G}). Finally, the relative spectrum of the action is the G representation $L_2(X, \mathcal{B}) \ominus L_2(X, \mathcal{G})$ (by which we mean the orthogonal complement to $L_2(X, \mathcal{G})$ in $L_2(X, \mathcal{B})$).

Note also that the notion of CPE makes sense also for cross-sections, since (relative) entropy is defined for them.

Proposition 4.6. *With the same assumptions as in Theorem 3.15, if the G action on X is completely positive entropy relative to \mathcal{G} , then S is completely positive entropy relative to \mathcal{G} .*

Proof. This follows from the proof of the second direction of Theorem 3.15 \square

Lemma 4.7. *There is a sequence of invariant sub-sigma-algebras $\dots \mathcal{G}_2 \subset \mathcal{G}_1 \subset \mathcal{G}_0 = \mathcal{B}$ such that the action of G on (X, \mathcal{G}_n) is CPE relative to \mathcal{G}_{n+1} , and $\cap \mathcal{G}_n$ is trivial (i.e. contains only null and conull sets).*

Proof. We construct the sequence by induction. Suppose \mathcal{G}_n has been constructed. Choose a partition P_n of X that is measurable with respect to \mathcal{G}_n , and such that $H(P) < \min\{h(G, \mathcal{G}_n), \frac{1}{n}\}$. Let \mathcal{F}_n be the invariant sigma-algebra generated by P . The factor $(X, \mathcal{B}) \rightarrow (X, \mathcal{F}_n)$ need not be CPE, but it has a Pinsker factor, which is the sigma-algebra generated by all partitions Q such that $h(Q|\mathcal{F}_n) = 0$. We take

\mathcal{G}_{n+1} be this sigma-algebra. The first requirement on \mathcal{G}_n is satisfied, and the second follows since if P is a partition measurable with respect to $\cap \mathcal{G}_n$ then

$$h(G, P) \leq h(G, \mathcal{G}_n) = h(G, \mathcal{F}_n) < \frac{1}{n}.$$

So $h(G, P) = 0$, and by the CPE assumption P is trivial. \square

For the second claim of the theorem, it is enough to show that $L_2(\mathcal{G}_n) \ominus L_2(\mathcal{G}_{n+1})$ is isomorphic to a countable sum of regular representations of G . We show that an analogue statement is also true for the first claim. First, recall the definition of the entropy of a partition, relative to a sigma-algebra.

Definition 4.8. Let $P = \{P_1, \dots, P_n\}$ be a partition of a probability space (X, \mathcal{B}, m) , and let $\mathcal{G} \subset \mathcal{B}$ be a sub-sigma-algebra. Let $p_i(x) = E(1_{P_i} | \mathcal{G})(x)$. Then for almost all x we have that $p_i(x) \geq 0$ and $\sum p_i(x) = 1$. We define

$$H(P | \mathcal{G}) = \int_X p_1(x) \log p_1(x) + \dots p_n(x) \log p_n(x) dx.$$

Lemma 4.9. *Suppose that the action of G on X is uniformly mixing relative to all \mathcal{G}_n 's. Then it is uniformly mixing.*

Proof. Let P be a partition and let $\epsilon > 0$. Choose n such that $H(P | \mathcal{G}_n) > H(P) - \epsilon$. By the assumption, there is a compact $K \subset G$ such that if $F \subset G$ is K -separated then

$$\frac{1}{|F|} H \left(\bigvee_{g \in F} gP \right) \geq \frac{1}{|F|} H \left(\bigvee_{g \in F} gP \middle| \mathcal{G}_n \right) > H(P | \mathcal{G}_n) - \epsilon > H(P) - 2\epsilon.$$

But the inequality in the other direction is trivial. \square

It is thus enough to show both relative claims for $n = 1$. Choose a cross-section S which is measurable with respect to \mathcal{G} and is U -discrete for some neighborhood of the identity in G . The cross-section $(S, \mathcal{B}, \mu, \mathcal{R}, \alpha : \mathcal{R} \rightarrow G)$ is such that α is \mathcal{G} -measurable. By Proposition 4.6 this cross-section is CPE relative to \mathcal{G} . By a theorem of [CFW] there is a transformation $T : S \rightarrow S$ that is \mathcal{G} -measurable and generates \mathcal{R} . This gives us another cross-section $(S, \mathcal{B}, \mu, \mathcal{R}, \beta : \mathcal{R} \rightarrow \mathbb{Z})$, and by Theorem 3.13 it is CPE relative to \mathcal{G} . Again, by Proposition 4.6 we conclude that T is CPE relative to \mathcal{G} , hence uniformly mixing. We wish to transfer this back to X .

Definition 4.10. We say that a cross-section $(S, \mathcal{B}, \mu, \mathcal{R}, \alpha)$ is uniformly mixing over a sub-sigma-algebra \mathcal{G} (such that α is \mathcal{G} -measurable) if, for all partitions P of S and $\epsilon > 0$, there is compact $K \subset G$ such that for any sequence $\phi_i : S \rightarrow S$, $i = 1, \dots, n$ that satisfies

1. $(x, \phi_i(x)) \in \mathcal{R}$ for all i and x .
2. $\alpha(x, \phi_i(x)) \notin K$.
3. ϕ_i is an isomorphism of measure spaces.

we have that

$$\left| \frac{1}{n} H \left(\bigvee_{i=1}^n \phi_i P \middle| \mathcal{G} \right) - H(P | \mathcal{G}) \right| < \epsilon.$$

In [RW] it is shown that CPE for a \mathbb{Z} action implies uniform mixing for the cross-section it generates. Moreover, it is shown there that (in our notation) the cross-section $(S, \mathcal{B}, \mu, \mathcal{R}, \alpha)$ is uniformly mixing over \mathcal{G} . It remains to show that this condition implies that the original G action is uniformly mixing.

Suppose P is a partition of X and $\epsilon > 0$. Choose a compact set $M \subset G$ such that $m(M \cdot S) > \epsilon$ and a neighborhood $V \subset G$ of the identity such that S is V -discrete and for any $g \in V$,

$$m(\{x | P(x) \neq P(gx)\}) < \epsilon.$$

Finally, choose a finite set $E \subset M$ such that $M \subset EV$. Define a partition Q on S to be the common refinement of gP where $g \in E$. Let $K \subset G$ be the compact set that one gets from the uniform relative mixing for the partition Q and for $\epsilon\lambda(V)$. We claim that if $F \subset G$ is KM -separated then the inequality (7) in Definition 4.2 holds. Assuming the contrary, since the sets $gV \cdot S$, $g \in E$ cover almost all X , there is $g \in E$ such that the fP , $f \in F$ are not $\epsilon\lambda$ independent on $gV \cdot S$. We can also assume that $m(g_i gV \cdot S \cap MS) \geq (1 - \epsilon)\lambda(V)$. Define $\phi_i : S \rightarrow S$ by $\phi_i(x) = hg_i gx$ if $g_i gx \in hS$ and $h \in M$ (and let $\phi_i(x) = x$ otherwise). Since the $g_i P$ are sufficiently dependent, we get that $\phi_i(Q)$ are also dependent. A contradiction.

We now move on to the spectral claim. We have that the action of T on S is CPE relative to \mathcal{G} . By Rohlin-Sinai there is a sigma-algebra \mathcal{H} on S such that $T\mathcal{H} \subset \mathcal{H}$, $\bigvee T^n \mathcal{H} = \mathcal{B}$, and $\bigwedge T^n \mathcal{H} = \mathcal{G}$. Choose an orthonormal basis ϕ_1, ϕ_2, \dots to $L_2(\mathcal{H}) \ominus L_2(T\mathcal{H})$. For any i , every neighborhood $V \subset G$ of the identity such that S is V^2 -discrete, and

every $g \in G$, define a function $\Phi_i^{g,V} \in L_2(X)$ by $\Phi_i^{g,V}(x) = \phi_i(y)$ if $y \in S$ and $x \in gV \cdot y$, and $\Phi_i^{g,V}(x) = 0$ otherwise. Note that if such y exists then it is unique.

Lemma 4.11. *Assume that S is W -discrete for some neighborhood of 1 in G . Suppose that $V, U \subset G$ are such that $V^2, U^2 \subset W$. Then*

1. *If $i \neq j$ then $\Phi_i^{g,V}$ and $\Phi_j^{h,U}$ are orthogonal.*
2. *$\langle \Phi_i^{g,V}, \Phi_i^{h,U} \rangle = \lambda(gV \cap hU)$.*

Moreover, the set $\{\Phi_i^{g,V}\}$ spans $L_2(\mathcal{B}) \ominus L_2(\mathcal{G})$.

Proof. We want to compute the inner product of $\Phi_i^{g,V}$ and $\Phi_j^{h,U}$. Without loss of generality we may assume $h = 1$. By definition, $\Phi_i^{g,V}(x)\Phi_j^{1,U}(x) = 0$ unless $x \in gVS$ and $x \in US$. Hence,

$$\langle \Phi_i^{g,V}, \Phi_j^{1,U} \rangle = \int_{s \in S} \int_{u \in U} \Phi_i^{g,V}(us) \phi_j(s) 1_{\{us \in gVS\}} du ds =$$

$$\int_S \int_{U \cap gV} \Phi_i^{g,V}(us) \phi_j(s) du ds + \int_S \int_{\{u \in U \setminus gV \mid us \in gVS\}} \Phi_i^{g,V}(us) \phi_j(s) du ds.$$

In the first integral, $\Phi_i^{g,V}(us) = \phi_i(s)$, so the integral is

$$\int_S \phi_i(s) \phi_j(s) ds \int_{U \cap gV} du = \langle \phi_i, \phi_j \rangle \lambda(U \cap gV).$$

As for the second integral, for each $s \in S$ and $u \in U$, the value of $\Phi_i^{g,V}(us)$ is either zero or equals $\phi_i(T^n s)$ for some $n = n(s, u) \in \mathbb{Z} \setminus \{0\}$. It is easy to see that n is independent of u , and that the function $s \rightarrow n(s)$ is \mathcal{G} -measurable. Moreover, the function R that assigns to s the measure of the set of $u \in U$ such that $us \in gVS$, is also \mathcal{G} -measurable. The second integral equals

$$\int_S \phi_i(s) \phi_j(T^{n(s)} s) R(s) ds = \sum_{k \neq 0} \int_S \phi_i(s) \phi_j(T^k s) R(s) 1_{\{n(s)=k\}} ds.$$

Suppose $k < 0$. Then $s \mapsto \phi_j(T^k s) R(s) 1_{\{n(s)=k\}}$ is $T^{-1}\mathcal{H}$ -measurable, and since $\phi_i \in L_2(\mathcal{H}) \ominus L_2(T^{-1}\mathcal{H})$, the integral is zero. If $k > 0$, the same argument works if we replace i and j . This proves 1. and 2. above.

Finally, assume $\Psi \in L_2(X)$. Then for almost all $g \in G$ we have that $\psi_g = g\Psi|_S$ is in $L_2(S)$. Suppose $\Psi \perp \Phi_i^{g,V}$ for all i, g and V . Then $\psi_g \perp T^n\phi_i$ for all i and n . Hence for almost all $g \in G$, we have $\psi_g \in L_2(S, \mathcal{G})$. Since \mathcal{G} is G -invariant, $\Psi \in L_2(X, \mathcal{G})$. \square

By the lemma, for every i , the map $1_{gV} \mapsto \Phi_i^{g,V}$ extends to a G -equivariant isometric embedding $L_2(G) \rightarrow L_2(X, \mathcal{B}) \ominus L_2(X, \mathcal{G})$, and the images of those maps for different i 's are orthogonal. By the second assertion of the lemma, these copies span. \square

References

- [Am] Ambrose, W.; *Representation of ergodic flows*. Ann. of Math., 42, 723–739.
- [A] Avni, N.; *Spectral and mixing properties of actions of amenable groups*. ERA 11 (2005) 57–63.
- [CFS] Cornfeld, I. P.; Fomin, S. V.; Sinai, Ya. G. *Ergodic theory*. Translated from the Russian by A. B. Sosinskiĭ. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], 245. Springer-Verlag, New York, 1982.
- [CFW] Connes, A.; Feldman, J.; Weiss, B. *An amenable equivalence relation is generated by a single transformation*. Ergodic Theory Dynamical Systems 1 (1981), no. 4, 431–450 (1982).
- [D] A. I. Danilenko *Entropy theory from orbital point of view*. Monatsh. Math., 134(2001), 121–141
- [DP] A. I. Danilenko, K. K. Park *Generators and Bernoullian factors for amenable actions and cocycles on their orbits*. Ergodic Theory Dynamical Systems 22 (2002), no. 6, 1715–1745.
- [DG] Dooley, A. H.; Golodets, V. Ya. *The spectrum of completely positive entropy actions of countable amenable groups*. J. Funct. Anal. 196 (2002), no. 1, 1–18.
- [FHM] Feldman, Jacob; Hahn, Peter; Moore, Calvin C. *Orbit structure and countable sections for actions of continuous groups*. Adv. in Math. 28 (1978), no. 3, 186–230.

- [K] Kamiński, B.; Liardet, P. *Spectrum of multidimensional dynamical systems with positive entropy*. Studia Math. 108 (1994), no. 1, 77–85.
- [L] Lindenstrauss, Elon *Pointwise theorems for amenable groups*. Invent. Math. 146 (2001), no. 2, 259–295.
- [OW] Ornstein, Donald S.; Weiss, Benjamin *Entropy and isomorphism theorems for actions of amenable groups*. J. Analyse Math. 48 (1987), 1–141.
- [OW2] Ornstein, Donald S.; Weiss, Benjamin *Limit theorems for amenable group actions* preprint.
- [RS] Rohlin, V. A.; Sinai, Ja. G. *The structure and properties of invariant measurable partitions. (Russian)* Dokl. Akad. Nauk SSSR 141 1961 1038–1041.
- [RW] Rudolph, Daniel J.; Weiss, Benjamin *Entropy and mixing for amenable group actions*. Ann. of Math. (2) 151 (2000), no. 3, 1119–1150.
- [W] B. Weiss *Actions of amenable groups*. In Topics in Dynamics and Ergodic Theory, edited by Sergey Bezuglyi and Sergiy Kolyada London Math. Soc. Lecture Note Series 310 (2003), 226–262.

Nir Avni
 Department of Mathematics
 Harvard University
 Cambridge, MA 02138
 avni.nir@gmail.com